Biochemical and Biophysical Research Communications 472 (2016) 237–242

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Crystal structure of the *Bacillus*-conserved MazG protein, a nucleotide pyrophosphohydrolase

Meong Il Kim, Minsun Hong*

Division of Biological Science and Technology, Yonsei University, Wonju 220-710, Republic of Korea

ARTICLE INFO

Article history:
Received 11 February 2016
Accepted 22 February 2016
Available online 23 February 2016

Keywords:
Bacillus anthracis
Bacillus cereus
MazG
Crystal structure
Nucleoside triphosphate pyrophosphohydrolase

ABSTRACT

BA1544 from *Bacillus anthracis* was previously annotated as a transcription factor for the gene cluster *ba1554* - *ba1558*, but has not been experimentally characterized. *B. anthracis* is an obligate pathogen causing fatal inhalational anthrax, and BA1544 is absolutely conserved in *Bacillus* species, including *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus mycoides*, with 100% sequence identity. To address the function of BA1544, we performed structural and biochemical studies, which revealed that BA1544 is a MazG protein. Thus, herein, the protein is defined as *Bacillus*-conserved MazG (BcMazG). Like other MazG structures, BcMazG assembles into a tetrameric architecture. Each monomer adopts a four-α-helix bundle that accommodates a metal ion using four acidic residues, and presents one putative substrate-binding site. Enzymatic characterization demonstrated that BcMazG is a nucleoside triphosphate (NTP) pyrophosphohydrolase and prefers adenosine triphosphate as a substrate among canonical NTPs. Moreover, structural comparison of BcMazG with its homologues revealed a potential regulation mechanism whereby the enzymatic activity of BcMazG is regulated by its C-terminal region.

1. Introduction

Bacillus anthracis is an obligate pathogen that causes fatal inhalational anthrax and can be used as a bioweapon. Thus, laboratory studies on *B. anthracis* are strictly regulated (class A agent, Centers for Disease Control, http://www.cdc.gov/). Unlike *B. anthracis*, *Bacillus cereus* is an opportunistic pathogen causing food poisoning that can be treated without extra medical care. Because *B. cereus* shares a large number of critical genes with *B. anthracis*, *B. cereus* proteins provide useful models for studying the function of their counterparts in *B. anthracis* [5].

Among various orthologous genes of *B. anthracis* and *B. cereus*, the gene cluster *ba1554* - *ba1558* from *B. anthracis* is highly conserved in gene organization and sequence with respect to the *bc1531* - *bc1535* cluster of *B. cereus* [1]. In particular, the first gene of the cluster, *ba1554*, shares 100% nucleotide sequence identity with *bc1531*. Furthermore, complete copies of *ba1554* have also been found in other *Bacillus* species, including *Bacillus thuringiensis* and *Bacillus mycoides*. Thus, as a *Bacillus*-conserved protein, BA1554 seems to play an indispensable role in *Bacillus* species.

BA1554 has been annotated as a putative transcription regulatory protein [1], presumably because the *ba1554* gene is located at the beginning of the *ba1554* - *ba1558* cluster. However, our sequence analysis suggests that BA1554 is homologous to *Escherichia coli* MazG (EcMazG), a pyrophosphohydrolase [19] with ~26% amino acid sequence identity. BA1554 also contains a MazG-like metal-coordinating motif that consists of four acidic residues, namely Glu37, Glu40, Glu66 and Asp69.

In the pFam protein family database, the MazG family (PF03819) is a large protein family, the members of which have diverse sequences. The MazG members contain a conserved canonical MazG domain of ~100 residues, characterized by an $EXXE_{12\text{-}28}EXXD$ motif (X, any amino acid residue; subscript number, the number of X residues). Numerous crystal structures of MazG family proteins have been determined and are available in the Protein Data Bank (PDB), revealing that each MazG domain folds into an α-helical bundle with a divalent magnesium cation that is coordinated by four acidic residues in a putative catalytic active site. In quaternary structures, MazG proteins assemble into homotetramers consisting of four identical MazG domains, or homodimers of two tandem MazG domains. Currently, our understanding of the MazG family is still limited because structural and biological studies have been reported for only a few MazG proteins, including three single-domain MazG proteins [integrin MazG (iMazG) from *Vibrio* sp.

* Corresponding author. 1 Yonseidae-gil, Mirae building #303, Wonju 200-710, Republic of Korea.
E-mail address: minsunhong@yonsei.ac.kr (M. Hong).

http://dx.doi.org/10.1016/j.bbrc.2016.02.097
0006-291X/

DAT722 [15], RS21-C6 from mouse [18], and dUTPase from *Deinococcus radiodurans* (DR2231) [4]] and two tandem MazG duplicates [EcMazG [6] and dUTPase from *Campylobacter jejuni* (CjdUTPase) [11]]. Biophysicochemical analyses of the MazG proteins have demonstrated that the MazG members are common in catalyzing a metal-dependent pyrophosphohydrolysis reaction, but have diverged to exhibit distinct substrate specificities and cellular functions. Among the MazG enzymes, the reported substrates differ, including non-specific nucleoside triphosphates (for EcMazG), dCTP/dATP (for iMazG) and dUTP (for dUTPases). Along with this wide range of substrate specificities, the MazG enzymes also have diverse physiological roles, such as the regulation of cell survival under nutritional or oxidative stresses and the removal of noncanonical NTPs, including 8-oxo-dGTP, dUTP and dITP, to prevent mutagenesis and DNA damage [4,6,9,12,15]

Here, we report structure-based biophysical analysis of *Bacillus*-conserved MazG (BcMazG). The crystal structure of BcMazG at 2.77 Å resolution reveals that BcMazG forms a dimer of homodimers through extensive hydrophobic interactions. Like other MazG enzymes, BcMazG catalyzes the pyrophosphohydrolysis of NTPs and preferentially converts ATP to AMP and PPi using Mg^{2+}. Structural comparison allowed us to locate a putative active site in close proximity to metal-coordinating residues. Furthermore, our structural modeling of a BcMazG-ATP complex and comparison with other MazG structures led us to propose a regulatory mechanism for BcMazG enzymatic activity.

2. Materials and methods

2.1. Production of the recombinant BcMazG protein

For the preparation of a bacterial expression construct of BcMazG, the open reading frame annotated as BC1531 (*i.e.*, BA1554) was amplified by PCR with the genomic DNA of *B. cereus* as a template. The PCR products with BamHI and SalI extensions were ligated into a pET49b vector that was modified to contain an N-terminal His_6 tag and thrombin cleavage site. The ligation mixture was transformed into DH5α. Sequences from the BcMazG expression construct were confirmed by DNA sequencing and used for the production of the recombinant BcMazG protein.

The BcMazG expression plasmids were re-transformed into a recombinant protein expression strain of *E. coli*, BL21 (DE3). Cells were grown at 37 °C in LB medium containing 50 μg/ml kanamycin. When the culture reached the exponential growth phase, protein production was induced by isopropyl β-D-1-thiogalactopyronoside for ~16 h at 18 °C.

Cells were pelleted by centrifugation. The cell paste in phosphate-buffered saline solution (PBS) and 10 mM imidazole was sonicated and cleared by centrifugation (~25,000 × g). The soluble BcMazG protein was purified by Ni-NTA (Qiagen) chromatography. The N-terminal tag was removed enzymatically by thrombin and the resulting tag-free BcMazG was purified by gel filtration chromatography on a Superdex 200 16/60 column (GE heathcare) in 20 mM HEPES (pH 7.4), 150 mM NaCl and 1.5 mM β-mercaptoethanol. The protein was concentrated to ~18.5 mg/mL for crystallization.

2.2. BcMazG crystallization and X-ray diffraction data collection

BcMazG crystals were obtained by the sitting-drop vapor-diffusion method in 0.1 mM sodium cacodylate pH 6.5 and 1 M sodium citrate. The crystal was cryoprotected by the supplementation of the crystallization solution with ethyleneglycol (30%), and was flash-frozen under a liquid nitrogen stream. X-ray diffraction data were collected at the Beamline 7A of the Pohang Accelerator Laboratory (PAL), autoindexed and merged by the Mosflm program [7], and scaled by the Scala program [3]. The summary of the crystallography statistics is shown in Table S1.

2.3. BcMazG structure determination

Initial phases for BcMazG were obtained by molecular replacement using the template of an unpublished YPJD structure (PDB ID CODE 2gta) with the Phaser program [10]. Structural models were visualized with COOT [2] and iteratively built for refinement. The structure was refined with the Refmac5 program [13]. The final structural model had good stereochemistry with no Ramachandran outliers. Structure refinement statistics are listed in Table S1.

2.4. Nucleoside triphosphate pyrophosphohydrolase assays

The NTP hydrolysis reaction was monitored by thin-layer chromatography (TLC) as described for the EcMazG protein [6]. In brief, NTPs, including ATP, GTP, CTP, and UTP to a final concentration of 1 mM, were added to a reaction solution containing 20 mM HEPES (pH 7.4), 150 mM NaCl and 2 mM $MgCl_2$ with BcMazG (0 μM, 20 μM, 200 μM). The reaction mixture was incubated at room temperature for 30 min for hydrolysis. The hydrolysis reaction was terminated by the addition of an equal volume of a stop solution containing 2% SDS and 20 mM EDTA. The final reactants were spotted on a polyethylenimine-cellulose TLC plate (Sigma), which was developed with a TLC running solution (0.75 M KH_2PO_4, pH 3.3). The TLC plates were visualized under a UV lamp so that nucleotide substrates or products could be observed.

NTP hydrolysis products were also analyzed via colorimetric assay, in which ammonium molybdate [16] reacts with inorganic phosphate (Pi) to form a blue compound. Hydrolysis reactions were performed in 200 μL of reaction buffer containing 20 mM HEPES (pH 7.4), 150 mM NaCl, 2 mM $MgCl_2$ and 5 μg of BcMazG with various amounts of ATP for 30 min at 24 °C, and were stopped by 15 min of heating at 70 °C. Then, the complete hydrolysis of pyrophosphate (PPi, from the BcMazG-ATP reaction) to Pi was catalyzed by the addition of an inorganic pyrophosphatase enzyme (Sigma). The amount of Pi was determined based on the optical density (OD) at a wavelength of 690 nm on a UV/visible spectrophotometer. For the determination of substrate specificities, NTP hydrolysis with each of the NTPs (ATP, GTP, CTP and UTP, 1 mM) was performed and the hydrolysis activity was normalized to that of ATP.

3. Results and discussion

3.1. The overall structure of BcMazG

Recombinant BcMazG was overexpressed in *E. coli* and purified to homogeneity in a soluble form (Fig. S1A). Then, BcMazG was crystallized and its crystal structure was determined to 2.77 Å resolution (Fig. 1). The asymmetric unit (ASU) contains four polypeptide chains of BcMazG (A, B, C and D) (Fig. 1C), which is consistent with an eluent with an apparent molecular weight of ~68 kDa in gel-filtration chromatography (Fig. S1B). The BcMazG structure contains residues 1-104 from the entire molecule (residues 1-113), and each subunit is conformationally similar, with root mean square deviation values of 0.18–0.24 Å for main-chain carbons. Chain A will mainly be described in this paper as a prototype structure of BcMazG unless specified.

The BcMazG tetramer consists of four helical bundles, similar to other MazG members. The BcMazG protomer contains four α-helices (H1-H4), their connecting loops (L1-L3), and the N- and C-terminal regions (NT and CT, respectively) in a topology of NT